BIER
Internet-Draft
Intended status: Informational
Expires: 1 September 2026

Z. Li
Z. Du
China Mobile
W. Cheng
J. Wang
G. Zhang
Centec Networks
28 February 2026

Scalable Data Plane Architecture for Bit Index Explicit Replication (BIER)

draft-li-bier-scalable-data-plane-00

Abstract

This document describes an optimized data plane forwarding architecture for Bit Index Explicit Replication (BIER). The standard BIER forwarding procedures, as conceptually described in RFC 8279, imply a per-bit evaluation of the BitString against the Bit Index Forwarding Table (BIFT) on each Bit Forwarding Router (BFR). When the BitStringLength is large (e.g., 256 bits or more), this conceptual model presents implementation challenges for high-speed ASIC/NPU forwarding engines in terms of processing latency and hardware table resources.

This document defines two conceptual data structures -- a Replication Member Table and an Interface BitMask Table -- that decouple the multicast replication decision from per-bit BitString parsing. The resulting forwarding pipeline operates in time proportional to the number of egress interfaces rather than the BitString length, enabling deterministic and scalable hardware processing. This mechanism is semantically equivalent to the forwarding procedures of RFC 8279 and does not modify the BIER encapsulation or introduce new protocol signaling.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Problem Statement 4
4. Architecture Overview 5
5. Conceptual Data Structures 7
  5.1. Replication Member Table (RMT) 7
  5.2. Interface BitMask Table (IBMT) 7
6. Forwarding Procedures 8
  6.1. Step 1: Replication Group Resolution 8
  6.2. Step 2: Interface-Level Replication 9
  6.3. Step 3: Egress BitString Isolation 9
  6.4. Step 4: Forwarding Decision 9
  6.5. Forwarding Example 9
7. Relationship to RFC 8279 Forwarding Procedures 11
8. Security Considerations 11
9. IANA Considerations 11
10. References 12
  10.1. Normative References 12
  10.2. Informative References 12
Acknowledgements 12
Authors' Addresses 12

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] provides a stateless multicast forwarding architecture that eliminates the need for per-flow state in transit routers and for explicit multicast tree-building protocols. In BIER, the ingress router (BFIR) encapsulates multicast data packets with a BIER header [RFC8296] that contains a BitString indicating the set of destination Bit-Forwarding Egress Routers (BFERs).

As BIER deployments scale to larger domains, multi-tenant data centers, and AI/ML cluster environments, the required BitStringLength (BSL) increases significantly. [RFC8279] Section 6.5 provides a conceptual forwarding pseudocode that iterates over each set bit in the BitString to determine the forwarding next-hop (BFR-NBR) and the per-neighbor Forwarding BitMask (F-BM). While RFC 8279 explicitly notes that this is a conceptual description and implementations may optimize the process, no standardized optimized architecture has been defined to guide hardware implementors.

This document fills this gap by defining a decoupled forwarding pipeline architecture that separates "determining which egress interfaces need packet copies" from "filtering the BitString for each copy." The key insight is that the number of distinct egress interfaces on a BFR is typically much smaller than the BitStringLength (e.g., 4-16 interfaces vs. 256+ bits), making interface-centric processing significantly more efficient for hardware implementation.

The mechanism described herein is semantically equivalent to the standard BIER forwarding procedures of RFC 8279 -- for any given input BitString, both approaches produce identical forwarding outcomes (same set of egress interfaces with the same per-interface BitStrings). This document does not modify the BIER encapsulation format defined in [RFC8296] and does not introduce any new protocol signaling.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

This document uses the following terms as defined in [RFC8279] and [RFC8296]:

BFR: Bit Forwarding Router.

BFIR: Bit-Forwarding Ingress Router.

BFER: Bit-Forwarding Egress Router.

BFR-NBR: A BFR that is a neighbor of the current BFR in the BIER forwarding topology.

BIFT: Bit Index Forwarding Table, as defined in [RFC8279].

BIFT-id: An identifier for a BIFT entry as defined in [RFC8296]. In MPLS encapsulation, this corresponds to an MPLS label value that implicitly encodes the <SD, SI, BSL> tuple.

BitString (BS): A string of bits in the BIER header indicating the set of destination BFERs.

F-BM (Forwarding BitMask): A bitmask associated with an egress interface, representing the logical OR of all BFR-id bit positions of downstream BFERs reachable via that interface.

This document introduces the following new terms:

Replication Group Identifier: A conceptual identifier derived from the BIER packet header’s BIFT-id, used to index the Replication Member Table. In practice, it MAY be identical to the BIFT-id.

Replication Member Table (RMT): A conceptual data structure that maps a Replication Group Identifier to a list of egress interface identifiers.

Interface BitMask Table (IBMT): A conceptual data structure that maps an egress interface identifier to the corresponding F-BM for that interface.

3. Problem Statement

The conceptual forwarding pseudocode in [RFC8279] Section 6.5 describes a process where each set bit in the incoming BitString is evaluated individually to determine the appropriate BFR-NBR and the associated F-BM. While this provides a clear and correct algorithmic specification, direct hardware implementation of this per-bit iteration model presents the following challenges in large-scale deployments:

First, processing latency scales with BitString length. When BSL is 256 or larger, scanning each bit position -- whether sequentially or in parallel -- imposes a processing cost that grows linearly with BSL. For line-rate forwarding of high-throughput multicast flows on 100GE or 400GE interfaces, this can become a significant constraint on forwarding engine design.

Second, hardware table resources scale with BitString length. A standard BIFT maintains one entry per BFR-id (i.e., per bit position), each containing a BFR-NBR identifier and a full-length F-BM. For BSL=256 and a single <SD, SI> combination, this requires 256 table entries, each storing a 256-bit F-BM, consuming substantial TCAM or SRAM resources. In resource-constrained edge routers or densely populated data center environments supporting multiple Sub-Domains and Set Identifiers, the aggregate table size can become prohibitive.

4. Architecture Overview

The optimized forwarding architecture decomposes BIER packet processing into two decoupled phases, each served by an independent conceptual lookup table. The following diagram illustrates the high-level forwarding pipeline:

```
                     BIER Packet Received
                              |
                              v
                  +-------------------+
                  | Extract BIFT-id   |
                  | from BIER Header  |
                  +--------+----------+
                           |
                           v
            +-------------------------------+
            | Phase 1: Replication Lookup  |
            | RMT[BIFT-id] -> {IF1,..,IFn} |
            +-------------------------------+
                           |
              +------------+------------+
              |            |            |
              v            v            v
         +--------+   +--------+   +--------+
         | Copy 1 |   | Copy 2 |   | Copy N |
         | -> IF1 |   | -> IF2 |   | -> IFN |
         +---+----+   +---+----+   +---+----+
             |            |            |
             v            v            v
      +-----------+ +-----------+ +-----------+
      | Phase 2:  | | Phase 2:  | | Phase 2:  |
      | BS AND    | | BS AND    | | BS AND    |
      | IBMT[IF1] | | IBMT[IF2] | | IBMT[IFN] |
      +-----------+ +-----------+ +-----------+
             |            |            |
             v            v            v
      +------------+ +------------+ +------------+
      | !=0: Fwd   | | !=0: Fwd   | | !=0: Fwd   |
      | ==0: Drop  | | ==0: Drop  | | ==0: Drop  |
      +------------+ +------------+ +------------+
```

Phase 1 (Replication Lookup) determines which egress interfaces require a copy of the packet, based solely on the BIFT-id. This is independent of the specific BitString content in the received packet.

Phase 2 (BitString Isolation) performs a single bitwise AND operation per egress interface to compute the filtered BitString for each copy. Copies with a resulting zero BitString are discarded.

The processing complexity of this pipeline is O(NumEgressInterfaces) for both phases combined, compared to O(BitStringLength) for the per-bit iteration model. In typical network topologies, a BFR has 2 to 16 egress interfaces, while BSL may be 256, 512, or larger.

5. Conceptual Data Structures

5.1. Replication Member Table (RMT)

A compliant BFR MUST maintain a conceptual Replication Member Table that maps each Replication Group Identifier to the set of egress interfaces that have at least one downstream BFER reachable through them.

Replication Group ID (derived from BIFT-id)	Egress Interface List
RG_1	{ IF_a, IF_b }
RG_2	{ IF_a, IF_c, IF_d }
...	...

The RMT is populated by the BIER control plane based on the network topology and BIFT computation results. The control plane MUST derive the egress interface list by identifying all unique BFR-NBRs present in the BIFT entries for the corresponding <SD, SI, BSL> and mapping each BFR-NBR to its associated egress interface.

The size of the RMT is O(N_groups), where N_groups is the number of distinct Replication Group Identifiers (i.e., distinct <SD, SI, BSL> tuples in use). Each entry contains a list of at most O(N_interfaces) egress interface identifiers.

5.2. Interface BitMask Table (IBMT)

A compliant BFR MUST maintain a conceptual Interface BitMask Table that maps each egress interface to the Forwarding BitMask (F-BM) representing all downstream BFERs reachable via that interface.

Egress Interface ID	Forwarding BitMask (F-BM)
IF_a	0b11000000...00 (BFR-ids 1,2 reachable)
IF_b	0b00110000...00 (BFR-ids 3,4 reachable)
IF_c	0b00001100...00 (BFR-ids 5,6 reachable)

For each egress interface IF_k, the F-BM is computed as the bitwise OR of all bit positions corresponding to BFERs whose shortest path (or engineered path, in the case of BIER-TE) from this BFR traverses IF_k. Formally:

F-BM(IF_k) = OR of { bit(BFR-id_j) } for all BFER_j where BIFT[BFR-id_j].BFR-NBR is reachable via IF_k.

The IBMT MUST be updated by the control plane whenever the BIER topology or BIFT entries change. The size of the IBMT is O(N_interfaces), where each entry stores one F-BM of BitStringLength bits.

Note: When the BFR participates in multiple Sub-Domains or supports multiple <SI, BSL> combinations, the IBMT entries MAY be scoped per <SD, BSL> or may use a single aggregate F-BM if the interface-to-downstream-BFER mapping is consistent across all applicable contexts. The specific scoping strategy is an implementation choice.

6. Forwarding Procedures

Upon receiving a valid BIER packet, a compliant BFR implementing this architecture MUST execute the following deterministic procedures in its data plane:

6.1. Step 1: Replication Group Resolution

The BFR extracts the BIFT-id from the received BIER packet header as specified in [RFC8296]. The BFR derives the Replication Group Identifier from the BIFT-id and queries the RMT to obtain the associated list of egress interface identifiers {IF_1, IF_2, ..., IF_n}.

If no matching entry exists in the RMT, the BFR MUST discard the packet and SHOULD log an error.

6.2. Step 2: Interface-Level Replication

The BFR generates an independent packet replica for each egress interface in the list obtained in Step 1. At this stage, the BFR MUST NOT modify the BitString in any of the replicas. All replicas carry the original BitString from the received packet.

6.3. Step 3: Egress BitString Isolation

For each packet replica destined to egress interface IF_k, the BFR queries the IBMT to retrieve the F-BM associated with IF_k. The BFR MUST compute:

BitString_new = BitString_original AND F-BM(IF_k)

6.4. Step 4: Forwarding Decision

For each replica:

* If BitString_new is zero (all bits clear), the BFR MUST discard the replica. This prevents forwarding packets to interfaces that have no relevant downstream BFERs for the current packet, avoiding unnecessary bandwidth consumption and potential forwarding anomalies.

* If BitString_new is non-zero, the BFR MUST overwrite the BIER header’s BitString field with BitString_new and transmit the replica via egress interface IF_k. The remaining BIER header fields (including BIFT-id) are updated according to the standard procedures of [RFC8279] and [RFC8296] (e.g., BIFT-id label swap in MPLS encapsulation, TTL decrement).

6.5. Forwarding Example

Consider BFR-A with three egress interfaces in a BIER domain using BSL=8 for simplicity of illustration:

```
                +---------+
                |  BFR-A  |
                +---------+
               /     |      \
            IF1     IF2     IF3
             |       |       |
          BFR-B   BFR-C   BFR-D
       (BFR-ids  (BFR-ids  (BFR-ids
         1, 2)    3, 4)     5, 6)
```

BFR-A’s RMT (for the relevant Replication Group):

Replication Group	Egress Interface List
RG_X	{ IF1, IF2, IF3 }

BFR-A’s IBMT:

Interface	F-BM
IF1	00000011
IF2	00001100
IF3	00110000

Scenario: BFR-A receives a BIER packet with BitString = 00101010 (destinations: BFR-ids 2, 4, 6).

```
Step 1: RMT[RG_X] -> {IF1, IF2, IF3}
Step 2: Create 3 replicas

Step 3 and 4:
  Replica 1 (IF1): 00101010 AND 00000011 = 00000010
                    Non-zero -> Forward with BS=00000010
  Replica 2 (IF2): 00101010 AND 00001100 = 00001000
                    Non-zero -> Forward with BS=00001000
  Replica 3 (IF3): 00101010 AND 00110000 = 00100000
                    Non-zero -> Forward with BS=00100000
```

Now consider BitString = 00101000 (destinations: BFR-ids 4, 6 only):

```
  Replica 1 (IF1): 00101000 AND 00000011 = 00000000
                    Zero -> DISCARD (no BFERs via IF1)
  Replica 2 (IF2): 00101000 AND 00001100 = 00001000
                    Non-zero -> Forward with BS=00001000
  Replica 3 (IF3): 00101000 AND 00110000 = 00100000
                    Non-zero -> Forward with BS=00100000
```

In the second scenario, the replica for IF1 is correctly discarded since no destination BFERs are reachable via IF1, demonstrating the isolation property of the architecture.

7. Relationship to RFC 8279 Forwarding Procedures

The forwarding mechanism described in this document is semantically equivalent to the conceptual forwarding procedures defined in [RFC8279] Section 6.5. Both approaches produce identical forwarding results for any given input BitString: the same set of egress interfaces receive packet copies, and each copy carries the same filtered BitString.

The difference lies in the iteration strategy. [RFC8279] iterates over bit positions (BFR-ids), grouping bits by their associated BFR-NBR. This document's approach pre-computes the bit-to-interface grouping into the RMT and IBMT, and iterates over egress interfaces instead. Since the number of egress interfaces (typically 2-16) is much smaller than the BitStringLength (typically 64-1024), this approach reduces the per-packet processing complexity from O(BSL) to O(N_interfaces).

A BFR implementing the architecture described in this document MUST produce forwarding results that are bit-for-bit identical to those produced by a correct implementation of [RFC8279] Section 6.5. Any deviation constitutes a forwarding error.

8. Security Considerations

This document does not introduce new security considerations beyond those already discussed in [RFC8279] and [RFC8296]. The forwarding architecture described herein is semantically equivalent to the standard BIER forwarding procedures and processes the same BIER header fields. No new protocol signaling or encapsulation formats are introduced.

The RMT and IBMT are populated by the BIER control plane and are subject to the same security considerations as the BIFT. Unauthorized modification of RMT or IBMT entries could result in incorrect packet replication or delivery to unintended interfaces. Implementations SHOULD protect these data structures with the same access control mechanisms used for BIFT entries.

9. IANA Considerations

This document has no IANA actions.

The mechanism defined in this document is a local implementation optimization for BIER forwarding. It does not define new protocol fields, encapsulation formats, or registry entries. The Replication Member Table and Interface BitMask Table are conceptual data structures internal to a BFR and do not require any protocol-level identifier allocation.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, 28 February 2026, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, 28 February 2026, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, 28 February 2026, <https://www.rfc-editor.org/info/rfc8279>.

[RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, 28 February 2026, <https://www.rfc-editor.org/info/rfc8296>.

10.2. Informative References

[RFC9262] Eckert, T., Ed., Menth, M., and G. Cauchie, "Tree Engineering for Bit Index Explicit Replication (BIER-TE)", RFC 9262, DOI 10.17487/RFC9262, 28 February 2026, <https://www.rfc-editor.org/info/rfc9262>.

Acknowledgements

The authors would like to thank the members of the BIER Working Group for their ongoing work on BIER scalability and deployment.

Authors’ Addresses

Zhiqiang Li
China Mobile
32 Xuanwumen West Street
Beijing
100053
China
Email: lizhiqiangyjy@chinamobile.com

Zongpeng Du
China Mobile
32 Xuanwumen West Street
Beijing
100053
China
Email: duzongpeng@chinamobile.com

Wei Cheng
Centec Networks
Suzhou
215000
China
Email: chengw@centec.com

Junjie Wang
Centec Networks
Suzhou
215000
China
Email: wangjj@centec.com

Guoying Zhang
Centec Networks
Suzhou
215000
China
Email: zhanggy@centec.com